A Randomized Evaluation of a Computer-Based Physician's Workstation: Design Considerations and Baseline Results

Barry L. Rotman[1,2], M.D., Andrea N. Sullivan[1], A.B., Thomas McDonald[1,2], M.D., Philippe DeSmedt[4], M.S., Donald Goodnature[4], M.S., Michael Higgins[4], Ph.D., Henri J. Suermondt[4], Ph.D., Charles Y. Young[4], Ph.D., Douglas K. Owens[1,2,3], M.D., M.Sc.

[1]Department of Veterans Affairs Medical Center, Palo Alto, CA [2]Division of General Internal Medicine, Stanford University, Stanford, CA [3]Department of Health Research and Policy, Stanford University, Stanford, CA [4]Hewlett-Packard Laboratories, Palo Alto, CA

ABSTRACT

We are performing a randomized, controlled trial of a Physician's Workstation (PWS), an ambulatory care information system, developed for use in the General Medical Clinic (GMC) of the Palo Alto VA. Goals for the project include selecting appropriate outcome variables and developing a statistically powerful experimental design with a limited number of subjects. As PWS provides real-time drug-ordering advice, we retrospectively examined drug costs and drug-drug interactions in order to select outcome variables sensitive to our short-term intervention as well as to estimate the statistical efficiency of alternative design possibilities. Drug cost data revealed the mean daily cost per physician per patient was 99.3¢ ± 13.4¢, with a range from 0.77¢ to 1.37¢. The rate of major interactions per prescription for each physician was 2.9% ± 1%, with a range from 1.5% to 4.8%. Based on these baseline analyses, we selected a two-period parallel design for the evaluation, which maximized statistical power while minimizing sources of bias.

1.0 BACKGROUND

A growing share of health care delivery occurs in an outpatient setting. The information needs for patient management in ambulatory practice have become increasingly complex as the nature of outpatient care has become more complicated [1]. Studies of clinicians in office practice have noted that according to practitioners, only a minority of their information needs are being met [2]. The medical informatics community has responded to this challenge. Information systems that originated in the hospital environment are expanding into the outpatient realm, and numerous outpatient systems are now commercially available. Although some randomized controlled trials of outpatient computer interventions have been conducted [3, 4], overall very few medical information systems have been evaluated [5] and even fewer have undergone a controlled trial [6]. We are conducting a randomized controlled trial in an ambulatory care clinic of a prototype Physician's Workstation (PWS) developed by Hewlett-Packard Laboratories.

Our trial assesses whether providing physicians with patient-specific information when they order outpatient medications can improve the quality and cost of medication therapy. Other researchers have documented drug cost reductions in the outpatient environment by pharmacist reviews of medications [7], computer-generated physician medication profiles [8], and computer-issued prescriptions based on personalized formularies [9]. Our project explores the area of reducing outpatient drug costs by online drug ordering. Tierney and colleagues have demonstrated a 15.3% reduction in inpatient drug costs with online drug ordering [10] . Our trial will also investigate the effect of real-time computer-based monitoring of drug-drug interactions. Although researchers have documented the frequency of drug interactions in ambulatory practice [11], there have been few attempts to document the effect of screening for interactions at the time of ordering.

In designing our evaluation, we confronted three challenges: (1) how to create a statistically viable study with a limited number of potential study subjects, (2) how to choose a study design that would control for potential confounding factors and secular trends, and (3) how to choose clinically significant health and economic outcomes that could be affected by a short-term intervention. We report here our approach to these challenges and the results of our baseline studies.

2.0 PHYSICIAN'S WORKSTATION

Researchers at Hewlett-Packard Laboratories have developed a prototype Physician's Workstation (PWS) to address the clinical information management problems in ambulatory care [12-14]. Functional goals for PWS include providing ready access to patient information stored in multiple files, presenting complex clinical data in ways that facilitate interpretation, and providing real-time clinical decision-making support while physicians use PWS for routine clinical tasks. PWS contains a graphical user interface that can present information

0195-4210/95/$5.00

from several databases, e.g. electrolyte levels and medications, in a single display, potentially providing the clinician with a more informed context for making decisions. PWS can graphically display laboratory data, demonstrating changes that occur over time. The system also includes a drug-ordering module, which contains decision support features that alert the physicians to patient-specific drug-drug, drug-disease, and drug-lab abnormality interactions when they prescribe medications. The drug-ordering module also provides drug costs, recommended cost-effective substitutions, and formulary restrictions when ordering prescriptions (see Table 1).

Table 1: Examples of alert messages for two drugs.

DRUG	MESSAGE
Terfenadine/ Seldane	Nonformulary. Please first try more cost-effective antihistamines (e.g., chlorpheniramine). Among non-sedating agents, astemizole (QD dosing) is least expensive.
Ciprofloxacin/ Cipro	Call ID for approval. Please consider cotrimoxazole, a more cost-effective antimicrobial agent.

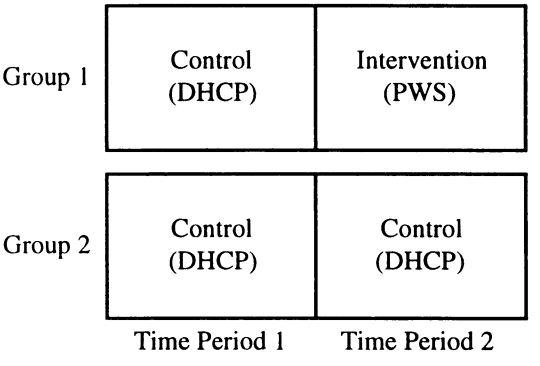
3.0 METHODS

3.1 Choice of Study Design

In considering alternative possibilities for the overall study design, we encountered several difficulties inherent in the evaluation of information systems. Because the PWS intervention was intended to influence the behavior of physicians (rather than patients), physicians are the appropriate unit of analysis. However, the number of potential study subjects consequently diminishes dramatically (from approximately 3,000 patients to approximately 35 physicians for our trial), which created the challenge of ensuring adequate statistical power for a trial. We therefore sought to select the experimental design that would maximize statistical power and minimize sources of bias.

We considered designs with historical controls, cross-over experiments, traditional parallel randomized designs with concurrent controls, and a two-period parallel randomized design. Experiments with historical controls are subject to bias introduced by secular trends. For example, as house officers' experience increases, their prescribing behavior may change independent of any experimental interventions. In cross-over designs, the investigator divides subjects into intervention and control groups that switch halfway through the trial. Although this design provides substantial statistical power even with small sample sizes, it is susceptible to bias due to the carry-over effect [15]. For example, an intervention occurring during one time period may exert an influence on a subsequent time period. We anticipated that our intervention would have a carry-over effect. For example, the physicians might learn and remember the prescribing advice, or a single prescription might have multiple refills that would spread the costs over both time periods. We therefore rejected the historically controlled and cross-over designs.

Figure 1: Two-Period Parallel Design



We then had to select between a traditional parallel randomized control design and a two-period parallel randomized design. A two-period parallel design consists of a trial with two time periods and two treatment groups [16, 17]. One group receives the control first and then the intervention, while the second group remains as the control during both treatment periods (see Figure 1). For each subject, the difference between outcomes in the two periods is calculated and the means of the differences between the treatment and control groups are compared using a two-sample t-test.

To choose between a traditional parallel design and a two-period parallel design, we estimated the statistical power of each design in our trial. The statistical power calculations for a traditional parallel design use total variance, σ^2_{total}, which is composed of the between-subject variance, σ^2_B (the variability in prescribed drug costs between physicians), and the within subject variance, σ^2_W (the variability in prescribed drug costs for a single physician over time):

$$\sigma^2_{total} = \sigma^2_B + \sigma^2_W \quad (1)$$

σ^2_{total} can be estimated directly from observable data, however, to calculate the power of the two-period parallel design requires an estimate of σ^2_W. We derived a convenient way to estimate σ^2_W by noting that the correlation between a physician's prescribed drug costs in two intervals, ρ, is related to σ^2_W:

$$\rho = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} \quad (2)$$

Since $\sigma^2_{total} = \sigma_B^2 + \sigma_W^2$, we can write:

$$\sigma_W^2 = \sigma^2_{total}(1-\rho) \quad (3)$$

We measured ρ in baseline studies as noted in section 4.1.

We performed power calculations for a two-sample t-test and assumed n = 16, a single-tailed α = 0.05, and power $(1-\beta)$ = 0.80. For the two-period parallel design, we used the within-subject variance and for the traditional parallel design, we used the total variance. We estimated the required effect size (change in costs of drugs prescribed by a physician) by using the following formulas [17, 18]:

Effect size (two-period parallel design)

$$= \sqrt{\frac{4\sigma_W^2(Z_\alpha + Z_\beta)^2}{n}} \quad (4)$$

Effect size (parallel design)

$$= \sqrt{\frac{2\sigma^2_{total}(Z_\alpha + Z_\beta)^2}{n}} \quad (5)$$

To collect the data required to estimate statistical power, we performed baseline studies of drug costs and drug-drug interactions. We performed a retrospective study of all prescriptions written for 2,890 GMC patients who were treated by 33 Stanford internal medicine house officers at the Department of Veterans Affairs Medical Center, Palo Alto (PAVAMC) from July 1, 1992, to June 30, 1993. The prescription data included drug name, quantity dispensed, unit price, dispensed dates, canceled dates, and prescribing physician and were obtained from the VA hospital information system, the Decentralized Hospital Computer Program (DHCP). The institutional review boards at Stanford University and the PAVAMC approved the study.

3.2 Drug Costs

Because our unit of analysis for the workstation intervention was the physician, we expressed our outcome variables in terms of each physician's drug costs. However, physicians cared for different numbers of patients. Therefore, we divided each physician's total prescription cost by the number of patients cared for by each physician. To account for patients being treated for different lengths of time we divided a patient's total costs by the number of days followed in clinic as determined by appointment and prescription data, yielding a daily cost per patient. Thus, we defined our outcome variable for drug cost as the mean daily cost per physician per patient.

Many patients had prescriptions that were not written by their General Medical Clinic (GMC) physicians, but rather by physicians in subspecialty clinics or the Emergency Department. For each physician's panel of patients we calculated the total drug costs, including prescriptions written by GMC as well as non-GMC physicians. We also calculated the fraction of total drug costs that were attributed to the GMC physician. For drug cost data, we calculated the mean and variance values, isolating total and within-subject variance by equations (1), (2), and (3).

3.3 Drug Interactions

To determine the prevalence of drug-drug interactions, we uploaded the prescriptions into the PWS system and analyzed them for interactions using the Drug Therapy Monitoring System (DTMS) by Medispan®. We created an algorithm that screened each patient's prescriptions in chronological order; thus, only drugs taken concurrently by a patient were screened for interactions. DTMS divides drug-drug interactions into five levels of severity, level 5 the least severe, level 1 the most severe. We analyzed only level 1 interactions, defined as an interaction with rapid onset, major severity, and established documentation. Many of the DTMS interactions, even level 1, were within the accepted standard of care but required careful monitoring. For example, a furosemide and digoxin interaction may cause no morbidity in the patient if the physician carefully monitors the patient's serum potassium level. By searching for evidence that the physician might not have been aware of a potential interaction, we defined a subset of interactions which have potential clinical relevance. A physician was considered unaware of an interaction if a relevant laboratory test was not performed within an appropriate time frame. For example, if a physician concurrently prescribed warfarin and ampicillin, generating a level 1 interaction, we examined the patient's laboratory data for evidence of a prothrombin time within 14 days of the interaction. In this case, we sought to determine if the physician was aware that the prescribed medications could cause an increased bleeding risk.

4.0 RESULTS

4.1 Drug Costs

The fraction of total drug costs for each GMC physician's panel of patients that were prescribed by the GMC physician varied from 20.7% to 61.1%, with a mean of 43.1% ± 10.0% (± standard deviation). The mean daily cost per physician per patient was 99.3¢ ± 13.4¢ for total drug costs, and varied from 0.77¢ to 1.37¢. To facilitate the statistical power calculations (see equations 4, 5), we analyzed the correlation between the cost of drugs that physicians prescribed during two consecutive four-month periods. The correlation coefficient between each physician's total prescribed costs during these time periods was 0.57.

4.2 Drug Interactions

Of the 20,723 prescriptions screened for potential drug-drug interactions, 616 generated a level 1 interaction in DTMS. The rate of interactions per prescription for each physician was 2.9% ± 1%, with a range from 1.5% to 4.8%. Digoxin was responsible for 64.3% of the level 1 interactions and warfarin caused 24.0%. Of these 616 prescriptions, 295 (47.8% of the level 1 interactions, or 1.4% of all prescriptions) were considered "clinically relevant" interactions; that is, these interactions were not monitored by the physician. Of the 295 clinically relevant interactions, warfarin accounted for 52.4%, and digoxin accounted for 50.3%. Analyzed by physician, the proportion of unmonitored interactions was 50% on average, and varied from 13.6% to 85.7%.

4.3 Choice of Study Design

We used the baseline studies on drug costs to determine the relative power of the two-period parallel design and the traditional parallel design. We therefore derived total and within-provider variance. For the total drug costs the sum of variances was σ^2_{total}=0.0180, within-subject variance was σ^2_W=0.0077, and between-subject variance was σ^2_B=0.0103. We estimated that the two-period parallel design would have an 80% chance of detecting a change in the mean drug cost per physician per day of 10.9¢ (the effect size), which was 0.8 standard deviation units or 11.0% of the total drug cost per physician per patient per day. The traditional parallel design was slightly less efficient, with an 80% chance of detecting a change of 11.8¢, which was 11.9% of mean daily costs or 0.9 standard deviation units. Thus, the two-period parallel design would detect smaller changes in physician behavior (measured as a decrease in the cost of drugs the physicians prescribed), and we chose it as the design for our trial.

5.0 DISCUSSION

The long-term goal of our study is to provide a rigorous evaluation of an innovative outpatient computer workstation. The workstation is designed to help physicians manage the increasingly complex information required in the care of outpatients in a general medical clinic. We found that the randomized two-period parallel design best satisfied our requirement for statistical efficiency and controlled for confounding and secular trends. We chose to examine the outcomes of drug cost and the number of drug-drug interactions because we believed these surrogate endpoints could be influenced by our intervention within the time frame of the trial, could be measured with the information available, and reflect potentially important health and economic outcomes.

Creating an operational definition of drug costs was difficult. In a practice environment with multiple providers prescribing medications for a patient, a potentially time consuming intervention aimed at one category of providers could lead to shifting of the prescription-writing burden to other non-intervention providers. For example, a GMC physician might choose to not use the drug ordering module in PWS and let the cardiologist renew a patient's antihypertensive medications. Our baseline data demonstrated a three-fold variation in the fraction of total drug costs written by each GMC physician. One possible explanation is that GMC physicians have differing opinions regarding the allocation of prescription writing. By selecting total drug costs as our outcome variable instead of drug costs prescribed by GMC physicians, we were able to eliminate variability associated with differences in allocation of prescribing responsibilities and to reduce variation in our outcome variable, which enhanced statistical power. Our effect size of a 10.9¢ reduction in daily cost per patient per physician approximates the estimate of a 9.6¢ reduction in daily drug costs that was generated at our institution by a study of manual medication list reviews and substitution recommendations by the pharmacy staff.

Our analysis of drug costs was subject to limitations in the prescription data, which were incomplete and did not reflect current pricing for all pharmaceuticals. We assumed, however, that these limitations were randomly distributed across all physicians. Thus, our estimates of total mean drug cost per day may underestimate total drug costs. The comparisons between physicians, however, are relatively accurate and are sufficient for experimental design calculations. The drug cost data will be updated prior to final analysis of the trial outcomes.

Our proportion of drug interactions per prescription, 2.9%, is within the range described in the literature. Jankel and Speedie reviewed the frequency of potential drug interactions in ambulatory patients and described a range of 1.2%-5.7% for major interactions [11]. We identified an important subset of interactions that lacked evidence that physicians were aware of interaction risk. The rates of unmonitored interactions represent an approximation of physician awareness. Our data does not attribute a laboratory study to a specific provider. Thus, if a provider other than the prescribing physician ordered an appropriate monitoring laboratory study, the prescribing physician was credited as being aware of the interaction, overestimating the rate of awareness. Conversely, a physician may have greater clinical understanding of a patient, for example, knowing that a patient has chronically been on the same doses of interacting medications and does not require laboratory

monitoring. In these instances our method would underestimate the rate of awareness. Despite these inaccuracies, the frequency of unmonitored interactions provides a more sensitive outcome variable than simply tallying all drug interactions, since many drug interactions are inherent in the practice of medicine. For example, 234 of our 616 interactions were digoxin and furosemide, a common combination in the treatment of congestive heart failure, which can cause toxicity if serum potassium falls too low. Thus, the overall frequency of drug interactions may be relatively resistant to modification. In contrast, the frequency of unmonitored interactions demonstrated greater than five-fold variation in our study. The frequency of unmonitored interactions, in particular 52.4% of coumadin interactions, suggests the potential benefit of screening for interactions at the time of drug ordering.

As computer systems move into the outpatient realm, evaluation methodologies need to adapt to the transition. We have designed a randomized controlled trial to assess the impact of a drug-ordering module within a an outpatient computer workstation. Our baseline data illustrates the need for cost variables that are viable in a multiprovider environment and for drug-interaction definitions that are sensitive to changes in physician behavior induced by a system such as PWS.

Acknowledgments: This study was supported by a grant from the VA Western Region Ambulatory Care Initiative. Dr. Rotman was supported by the VA Ambulatory Care Fellowship program. Dr. Owens is supported by a Career Development Award from the VA Health Services Research and Development Service. We thank the staff of Information Resources Management Service (IRMS) VA Medical Center, Palo Alto, for their support. Our thanks to Debby Fife for comments on the manuscript, and to Byron W. Brown, PhD for advice on experimental design.

References

1. McDonald C and Tierney W. Computer-stored medical records. Their future role in medical practice. *JAMA*. 1988;259:3433-40.
2. Covell D, Uman G and Manning P. Information needs in office practice: Are they being met? *Ann Intern Med*. 1985;103:596-9.
3. McPhee S, Bird J, Fordham D, Rodninck J and Osborn E. Promoting cancer prevention activities by primary care physicians: Results of a randomized, controlled trial. *JAMA*. 1991;266:538-44.
4. McDonald C, Hui S, Smith D, et al. Reminders to physicians from an introspective computer medical record: A two-year randomized trial. *Ann Intern Med*. 1984;100:130-8.
5. Tierney W, Overhage J and McDonald C. A plea for controlled trials in medical informatics. *JAMA*. 1994;1:353-5.
6. Balas E, Austin S, Brown G and Mitchell J. Quality evaluation of controlled clinical information service trials. *Proceedings of the Seventeenth Annual Symposium on Computer Applications in Medical Care*; Washington, DC, 1993:586-90.
7. Britton M and Lurvey P. Impact of medication profile review on prescribing in a general medicine clinic. *Am J Hosp Pharm*. 1991;48:265-70.
8. Hershey C, Porter D, Breslau D and Cohen D. Influence of simple computerized feedback on prescription charges in an ambulatory clinic. *Med Care*. 1986;24:472-81.
9. Donald J. Prescribing costs when computers are used to issue all prescriptions. *BMJ*. 1989;299:28-30.
10. Tierney W, Miller M, Overhage J and McDonald C. Physician inpatient order writing on microcomputer workstations. *JAMA*. 1993;269:379-83.
11. Jankel C and Speedie S. Detecting drug interactions: A review of the literature. *DICP*. 1990;24:982-9.
12. Tang P, Annevelink J, Fafchamps D, Stanton W and Young C. Physician's Workstations: integrated information management for clinicians. *Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care*; Washington, DC, 1991:569-73.
13. Tang P, Annevelink J, Suermondt H and Young C. Semantic integration of information in a physician's workstation. *Int J BioMed Comput*. 1993;35:47-60.
14. De Smedt P, Annevelink J, Pham T and Strong P. A Physician's Workstation as an Application of Object-Oriented Database Technology in Healthcare. *Proceedings of the 1994 International Conference on Applications of Databases (ADB-94)*; Vadstena, Sweden, 1994
15. Brown B. The crossover experiment for clinical trials. *Biometrics*. 1980;36:69-79.
16. Kramer M and Shapiro S. Scientific challenges in the application of randomized trials. *JAMA*. 1984;252:2739-45.
17. Chassan J. A note on relative efficiency in clinical trials. *J Clin Pharmacol*. 1970;10:359-60.
18. Ott L. *An introduction to statistical methods and data analysis*. Third ed. Boston: PWS-Kent Publishing company; 1988.